

From: Chan, Christina
Sent: Tuesday, May 31, 2005 5:32 PM
To: Bausch, Sarae; STIC-Biotech/ChemLib
Subject: RE: sequence rush request 10/009340

CRFB

Please rush. Thanks Chris

10-009340

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Bausch, Sarae
Sent: Tuesday, May 31, 2005 1:01 PM
To: Chan, Christina
Subject: sequence rush request 10/009340

Could I get a rush nucleic acid sequence search for SEQ ID No. 1 (932 nt)? Please print out the first 100 hits.

Thank you.

Sarae Bausch, Ph.D.
USPTO Art Unit 1634
REM 2 E 84
Mailbox: REM 2 C 70
(571) 272-2912

1-932NA
LB*****
STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable.

STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

Date completed: 6-3
Searcher: Beverly e 2528
Terminal time: _____
Elapsed time: _____
CPU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Search Site

_____ STIC
_____ CM-1
_____ Pre-S

Type of Search

_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors

_____ IG
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
✓ _____ Other CGN

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 16:52:41 ; Search time 203 Seconds

(without alignments)
7512.364 Million cell updates/sec

Title: US-10-009-340-1

Perfect score: 932

Sequence: 1 cagctgtgactgtgtat.....agagcaaacactggcacag 932

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	82.4	8.8	700	4	US-09-202-329-1
C 2	78.6	8.4	2182	4	US-09-669-980A-8
C 3	77.4	8.3	365	4	US-09-621-976-12759
C 4	77	8.3	1441	4	US-09-220-132-39
C 5	75.6	8.1	960	4	US-09-248-796A-6497
C 6	75.6	8.1	1888	1	US-08-229-145-13
C 7	75	8.0	1594	2	US-08-935-450-3
C 8	75	8.0	1594	4	US-09-338-123-3
C 9	75	8.0	38239	4	US-09-949-016-12348
C 10	75	8.0	38252	4	US-09-949-016-13570
C 11	74	7.9	52226	4	US-09-949-016-14426
C 12	74	7.9	63644	4	US-09-949-016-12098
C 13	72.8	7.8	305	3	US-09-118-554-43
C 14	72.8	7.8	305	3	US-09-118-627-43
C 15	72.8	7.8	305	3	US-09-602-877A-43
C 16	66	7.1	421	3	US-09-404-879A-120
C 17	66	7.1	421	4	US-09-338-933-120
C 18	66	7.1	421	4	US-09-215-681-120
C 19	66	7.1	421	4	US-09-216-003A-120
C 20	66	7.1	421	4	US-09-667-857-120
C 21	63.8	6.8	2295	4	US-09-774-528-252
C 22	63.8	6.8	2709	4	US-09-774-528-251
C 23	61.2	6.6	76472	4	US-09-949-016-15896
C 24	61.2	6.6	92276	4	US-09-949-016-12166
C 25	56.2	6.0	1141	4	US-09-806-708B-22
C 26	54	5.8	1141	4	US-09-806-708B-22
C 27	48.6	5.2	18989	4	US-09-949-016-13435

28	48.6	5.2	21885	4	US-09-949-016-12600	Sequence 12600, A
C 29	48.2	5.2	615	3	US-08-998-416-186	Sequence 186, App
C 30	44.4	4.8	168575	3	US-09-426-290-1	Sequence 1, Appli
C 31	44	4.7	5340	4	US-09-627-122-21	Sequence 21, Appl
C 32	43.4	4.7	17865	4	US-09-949-016-15950	Sequence 15950, A
C 33	43.2	4.6	55886	4	US-09-949-016-15129	Sequence 15129, A
C 34	43.2	4.6	156324	4	US-09-949-016-13749	Sequence 13749, A
C 35	43	4.6	731	1	US-08-451-405A-2	Sequence 2, Appli
C 36	42.6	4.6	1055	4	US-09-806-708B-23	Sequence 23, Appli
C 37	42.4	4.5	7620	4	US-09-949-016-16939	Sequence 16939, A
C 38	42.4	4.5	95255	4	US-09-949-016-17067	Sequence 17067, A
C 39	42.2	4.5	7218	1	US-08-232-463-14	Sequence 14, Appl
C 40	42.2	4.5	264665	4	US-09-949-016-13747	Sequence 13747, A
C 41	42	4.5	16421	4	US-09-949-016-16169	Sequence 16169, A
C 42	41.4	4.4	1182	4	US-09-601-198-18	Sequence 18, Appl
C 43	41.4	4.4	3095	6	5231168-1	Patent No. 5231168
C 44	41.4	4.4	3095	6	5231168-1	Patent No. 5231168
C 45	41.4	4.4	18773	4	US-09-949-016-14164	Sequence 14164, A
C 46	41.2	4.4	832	4	US-09-621-976-2813	Sequence 2813, App
C 47	41	4.4	601	4	US-08-949-016-205148	Sequence 205148, A
C 48	41	4.4	642	1	US-08-764-100-13	Sequence 13, Appli
C 49	41	4.4	3000	1	US-08-764-100-9	Sequence 9, Appli
C 50	41	4.4	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 51	41	4.4	218940	4	US-09-949-016-17539	Sequence 17539, A
C 52	41	4.4	640681	4	US-09-790-988-1	Sequence 1, Appli
C 53	40.8	4.4	1939	1	US-07-715-751B-2	Sequence 2, Appli
C 54	40.6	4.4	253345	4	US-09-949-016-12656	Sequence 12656, A
C 55	40.6	4.4	253364	4	US-09-949-016-13639	Sequence 13639, A
C 56	40.6	4.4	343352	4	US-09-949-016-13498	Sequence 13498, A
C 57	40.4	4.3	636	3	US-08-998-416-1137	Sequence 1137, App
C 58	40.4	4.3	1137	3	US-09-540-236-763	Sequence 763, App
C 59	40.4	4.3	5923	3	US-08-064-922-3	Sequence 3, Appli
C 60	40.4	4.3	94750	4	US-09-596-002-38	Sequence 38, Appli
C 61	40.2	4.3	640681	4	US-09-790-988-1	Sequence 1, Appli
C 62	40	4.3	43896	4	US-09-949-016-14942	Sequence 14942, A
C 63	40	4.3	67755	4	US-09-949-016-13703	Sequence 13703, A
C 64	39.8	4.3	601	4	US-09-949-016-144389	Sequence 144389, A
C 65	39.8	4.3	601	4	US-09-949-016-144390	Sequence 144390, A
C 66	39.8	4.3	601	4	US-09-949-016-144391	Sequence 144391, A
C 67	39.8	4.3	1548	4	US-09-270-767-30207	Sequence 30207, A
C 68	39.8	4.3	2666	4	US-09-270-767-14105	Sequence 14105, A
C 69	39.8	4.3	29357	4	US-09-949-016-16676	Sequence 16676, A
C 70	39.8	4.3	55885	4	US-09-949-016-15129	Sequence 15129, A
C 71	39.8	4.3	133613	4	US-09-949-016-15824	Sequence 15824, A
C 72	39.6	4.2	615	3	US-08-998-416-186	Sequence 186, App
C 73	39.6	4.2	636	3	US-08-998-416-1137	Sequence 1137, App
C 74	39.6	4.2	837	3	US-08-998-416-288	Sequence 288, App
C 75	39.6	4.2	1388	1	US-08-282-581-3	Sequence 3, Appli
C 76	39.6	4.2	1388	1	US-08-550-544-3	Sequence 3, Appli
C 77	39.4	4.2	456	4	US-09-270-767-29649	Sequence 29649, A
C 78	39.4	4.2	759	4	US-09-270-767-13637	Sequence 13637, A
C 79	39.4	4.2	2007	4	US-09-270-767-10846	Sequence 10846, A
C 80	39.4	4.2	6996	4	US-09-573-080A-29	Sequence 29, Appl
C 81	39.2	4.2	735	4	US-09-710-279-2779	Sequence 2779, App
C 82	39.2	4.2	747	3	US-09-134-001C-2248	Sequence 2248, App
C 83	39.2	4.2	3001	4	US-09-539-333D-185	Sequence 185, App
C 84	39.2	4.2	4136	4	US-09-710-279-3686	Sequence 3686, App
C 85	39.2	4.2	181429	4	US-09-949-016-12372	Sequence 12372, A
C 86	39.2	4.2	181430	4	US-09-949-016-15772	Sequence 15772, A
C 87	39	4.2	2358	4	US-09-248-796A-10389	Sequence 10389, A
C 88	39	4.2	25951	4	US-09-949-016-16194	Sequence 16194, A
C 89	39	4.2	236341	4	US-09-949-016-13978	Sequence 13978, A
C 90	39	4.2	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 91	39	4.2	1664976	4	US-09-692-570-1	Sequence 1, Appli
C 92	38.8	4.2	55827	3	US-09-813-133A-3	Sequence 3, Appli
C 93	38.8	4.2	55827	4	US-10-212-877-3	Sequence 3, Appli
C 94	38.8	4.2	59519	4	US-09-949-016-13504	Sequence 13504, A
C 95	38.8	4.2	137226	4	US-09-949-016-13763	Sequence 13763, A
C 96	38.6	4.1	601	4	US-09-949-001-580	Sequence 580, App
C 97	38.6	4.1	1093	4	US-09-270-767-10614	Sequence 10614, A
C 98	38.6	4.1	1174	4	US-09-270-767-10951	Sequence 10951, A
C 99	38.6	4.1	3435	4	US-09-710-279-3648	Sequence 3648, App
C 100	38.6	4.1	27587	4	US-09-949-016-12185	Sequence 12185, A

ALIGNMENTS

```

RESULT 1
US-09-202-329-1/c
/ Sequence 1, Application US/09202329A
/ Patent No. 6676944
/ GENERAL INFORMATION:
/ APPLICANT: Dalton, John P
/ APPLICANT: Andrews, Stuart J
/ TITLE OF INVENTION: Vaccine containing a peroxiredoxin and/or a B-tubulin
/ FILE REFERENCE: 1181-243
/ CURRENT APPLICATION NUMBER: US/09/202,329A
/ CURRENT FILING DATE: 1999-02-19
/ EARLIER APPLICATION NUMBER: GB 9612214.8
/ EARLIER FILING DATE: 1996-06-11
/ EARLIER APPLICATION NUMBER: PCT/GB97/01573
/ EARLIER FILING DATE: 1997-06-11
/ NUMBER OF SEQ ID NOS: 21
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 700
/ TYPE: DNA
/ ORGANISM: Fasciola hepatica
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (2)..(700)
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (5)
/ OTHER INFORMATION: n = a, or t, or c, or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (9)
/ OTHER INFORMATION: n=a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (11)
/ OTHER INFORMATION: n= a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (16)
/ OTHER INFORMATION: n= a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (18)
/ OTHER INFORMATION: n= a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (26)
/ OTHER INFORMATION: n= a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (42)
/ OTHER INFORMATION: n= a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (44)
/ OTHER INFORMATION: n= a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (55)
/ OTHER INFORMATION: n = a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (57)
/ OTHER INFORMATION: n= a or t or c or g
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (62)
/ OTHER INFORMATION: n= a or t or c or g

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 16:02:00 ; Search time 3731 seconds
(without alignment)
9508.416 Million cell updates/sec

Title: US-10-009-340-1
Perfect score: 932
Sequence: 1 cagctgtagcatcttgatat.....agagcaacaacctggcacag 932

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsl1:*

9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	372	39.9	372	8	B2662094
C 2	189.6	20.3	247	8	BH856235
C 3	186	20.0	473	8	CC460355
C 4	107.2	11.5	509	7	CK851987
C 5	107.2	11.5	546	5	BQ047631
C 6	107.2	11.5	564	7	CV430431
C 7	107.2	11.5	631	7	CV431748
C 8	107.2	11.5	656	7	CV502060
C 9	107.2	11.5	736	7	CK267151
C 10	107.2	11.5	755	7	CK260303
C 11	107.2	11.5	783	7	CK214137
C 12	107.2	11.5	792	7	CK266623
C 13	107.2	11.5	827	7	CK260306
C 14	107.2	11.5	840	7	CV429559
C 15	107.2	11.5	886	4	BG129116
C 16	105.6	11.3	654	4	BG129116
C 17	105.6	11.3	694	7	CK273826
C 18	104	11.2	367	5	BP631495
C 19	104	11.2	386	1	AV802029
C 20	104	11.2	403	1	AV818452
C 21	104	11.2	407	1	AV804835
C 22	104	11.2	409	5	BP639156
C 23	104	11.2	411	5	BP668549
C 24	104	11.2	418	5	BP607321

25	104	11.2	419	5	BP608845
26	104	11.2	426	1	AV797046
27	104	11.2	426	1	AV797046
28	104	11.2	427	5	BP627021
29	104	11.2	437	5	BP593288
30	104	11.2	440	5	BP610928
31	104	11.2	444	1	AV562504
32	104	11.2	450	5	BP65295
33	104	11.2	568	1	AV561222
34	104	11.2	638	8	AQ959863
35	104	11.2	665	1	AV782583
36	104	11.2	699	4	BM111635
37	104	11.2	840	7	CK270157
38	104	11.2	858	7	CK274174
39	104	11.2	863	7	CK270156
40	103	11.1	437	1	AV787752
41	103	11.1	441	5	BP633999
42	103	11.1	451	1	AV794292
43	103	11.1	532	7	CN521712
44	103	11.1	564	7	CV227058
45	102.4	11.0	381	5	BP634984
46	102.4	11.0	409	5	BP651750
47	102.4	11.0	411	5	BP638315
48	102.4	11.0	439	1	AV796223
49	102.4	11.0	448	5	BP589971
50	102.4	11.0	451	5	BP589982
51	102.4	11.0	452	5	BP590731
52	101.4	10.9	462	5	BP593870
53	101.4	10.9	541	7	CN517269
54	101.4	10.9	565	7	CV255862
55	101.4	10.9	625	1	AJ778117
56	101.4	10.9	675	7	CV233458
57	101.4	10.9	699	7	CV249367
58	101.4	10.9	705	7	CV250156
59	101.4	10.9	707	7	CV253314
60	101.4	10.9	773	7	CV260213
61	101.4	10.9	790	7	CN523196
62	101.4	10.9	923	7	CV247172
63	100.8	10.8	421	5	BP614006
64	99.8	10.7	606	7	CK087314
65	99.8	10.7	951	7	CV429463
66	99.6	10.7	429	5	BP606897
67	99.2	10.6	662	7	CN744157
68	98.6	10.6	290	1	AV543197
69	98.6	10.6	400	1	AV816610
70	98.6	10.6	400	5	BP573873
71	98.6	10.6	401	1	AV801224
72	98.6	10.6	408	1	AV805657
73	98.6	10.6	409	7	Z34951
74	98.6	10.6	409	7	Z35206
75	98.6	10.6	409	7	Z35363
76	98.6	10.6	410	1	AV801176
77	98.6	10.6	411	5	BP588995
78	98.6	10.6	413	1	AV790322
79	98.6	10.6	414	1	AV803577
80	98.6	10.6	417	1	AV811393
81	98.6	10.6	418	5	BP607519
82	98.6	10.6	419	1	AV804137
83	98.6	10.6	420	1	AV810744
84	98.6	10.6	421	1	AV798422
85	98.6	10.6	422	1	AV820641
86	98.6	10.6	426	1	AV564564
87	98.6	10.6	429	1	AV796641
88	98.6	10.6	430	5	BP612995
89	98.6	10.6	434	5	BP608266
90	98.6	10.6	438	5	BP609740
91	98.6	10.6	441	5	BP586386
92	98.6	10.6	449	5	BP590642
93	98.6	10.6	451	5	BP585346
94	98.6	10.6	452	5	BP590510
95	98.6	10.6	475	1	AV565801
96	98.6	10.6	479	1	AV566503
97	98.6	10.6	489	1	AV560529

BP608845	BP608845
AV797046	AV797046
BP627021	BP627021
BP593288	BP593288
BP610928	BP610928
AV562504	AV562504
BP65295	BP65295
AV561222	AV561222
AQ959863	AQ959863
AV782583	AV782583
BM111635	BM111635
CK270157	CK270157
CK274174	CK274174
CK270156	CK270156
AV787752	AV787752
BP633999	BP633999
AV794292	AV794292
CN521712	CN521712
CV227058	CV227058
BP634984	BP634984
BP651750	BP651750
BP638315	BP638315
AV796223	AV796223
BP589971	BP589971
BP589982	BP589982
BP590731	BP590731
BP593870	BP593870
CN517269	CN517269
CV255862	CV255862
AJ778117	AJ778117
CV233458	CV233458
CV249367	CV249367
CV250156	CV250156
CV253314	CV253314
CV260213	CV260213
CN523196	CN523196
CV247172	CV247172
BP614006	BP614006
CK087314	CK087314
CV429463	CV429463
BP606897	BP606897
CN744157	CN744157
AV543197	AV543197
AV816610	AV816610
BP573873	BP573873
AV801224	AV801224
AV805657	AV805657
Z34951	Z34951
Z35206	Z35206
Z35363	Z35363
AV801176	AV801176
BP588995	BP588995
AV790322	AV790322
AV803577	AV803577
AV811393	AV811393
BP607519	BP607519
AV804137	AV804137
AV810744	AV810744
AV798422	AV798422
AV820641	AV820641
AV564564	AV564564
AV796641	AV796641
BP612995	BP612995
BP608266	BP608266
BP609740	BP609740
BP586386	BP586386
BP590642	BP590642
BP585346	BP585346
BP590510	BP590510
AV565801	AV565801
AV566503	AV566503
AV560529	AV560529

```

98 98.6 10.6 493 1 AV566036 AV566036 linear GSS 31-JAN-2003
99 98.6 10.6 506 1 AV548335 AV548335
100 98.6 10.6 544 1 AV558439 AV558439

ALIGNMENTS

RESULT 1
BZ662094/c
LOCUS BZ662094 372 bp DNA linear GSS 31-JAN-2003
DEFINITION SALK_025578.41.40.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_025578.41.40.x, genomic
survey sequence.
ACCESSION BZ662094
VERSION BZ662094.1 GI:28176052
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 372)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
Location/Qualifiers
1..372
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_025578.41.40.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

FEATURES
source
1..372
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_025578.41.40.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
Query Match 39.9%; Score 372; DB 8; Length 372;
Best Local Similarity 100.0%; Pred. No. 8.1e-71;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 AGCTTATCTCAACATCAGCGTGATATAACCAATATCTTACTTCTTTTACATTTGTGAAT 172
Db 372 AGCTTATCTCAACATCAGCGTGATATAACCAATATCTTACTTCTTTTACATTTGTGAAT 313

QY 173 GGAACCAACCCATTTTCTGGAAAAGTGCTAACCAACATTTGATTACCGTATCACTA 232
Db 312 GGAACCAACCCATTTTCTGGAAAAGTGCTAACCAACATTTGATTACCGTATCACTA 253

QY 233 CTACCTTTCATTTCTATCTTCTGTTTCATTATGCTGACTATTTTAAGCTCCGTTGTCAAATC 292
Db 252 CTACCTTTCATTTCTATCTTCTGTTTCATTATGCTGACTATTTTAAGCTCCGTTGTCAAATC 193

QY 293 TCTAAGTTAGACATAAAAGACAAAGACTAATCAATTTGTTCATCACACCGCGCTGTCGAGT 352

```

```

Db 192 TCTAAGTTAGACATAAAAGACAAAGACTAATCAATTTGTTCATCACACCGCGCTGTCGAGT 133
QY 353 GAGCTATATTAACTCGTGGATTTTAAAGCATTAAAGAAACATTTCTATAGTACTTAAAGCAAAT 412
Db 132 GAGCTATATTAACTCGTGGATTTTAAAGCATTAAAGAAACATTTCTATAGTACTTAAAGCAAAT 73
QY 413 AAAATAATTATTAATCAAAACACTATGCTTGACACTGGTGCAGTGCTAGTGGTAGTGAATGAT 472
Db 72 AAAATAATTATTAATCAAAACACTATGCTTGACACTGGTGCAGTGCTAGTGGTAGTGAATGAT 13
QY 473 TCTACATCAATAA 484
Db 12 TCTACATCAATAA 1

RESULT 2
BH856235/c
LOCUS BH856235 247 bp DNA linear GSS 08-JUL-2002
DEFINITION SALK_083148.19.45.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_083148.19.45.x, genomic
survey sequence.
ACCESSION BH856235
VERSION BH856235.1 GI:21705825
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 247)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
At2g34760.
Class: TDNA tagged.
Location/Qualifiers
1..247
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_083148.19.45.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

FEATURES
source
1..247
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_083148.19.45.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
Query Match 20.3%; Score 189.6; DB 8; Length 247;
Best Local Similarity 87.9%; Pred. No. 5.5e-31;
Matches 218; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

QY 117 TATCTCAACATCAGCGTGATATAACCAATATCTTACTTCTTTTACATTTGTGAATGAA 176
Db 247 TATCTCAACATCAGCGTGATATAACCAATATCTTACTTCTTACAGGTGTGAATGAA 189

QY 177 CCAACCCCATTTTCTGGAAAAGTGCTAACCAACATTTGATTACCGTATCACTACTAC 236

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 17:59:13 ; Search time 676 Seconds
(without alignments)
8475.441 Million cell updates/sec

Title: US-10-009-340-1

Perfect score: 932

Sequence: 1 cagctgtgacattgtgat.....agagcaaacctggcacag 932

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706592 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	104	11.2	1335	9	US-09-938-842A-622
C 2	104	11.2	1335	11	US-09-938-842A-622
C 3	104	11.2	1469	9	US-09-770-445-8
C 4	100.8	10.8	773	17	US-10-333-184-431
C 5	99.2	10.6	469	17	US-10-333-184-263
C 6	98.6	10.6	445	9	US-09-770-444-806
C 7	98.6	10.6	586	17	US-10-333-184-303
C 8	98.6	10.6	641	9	US-09-910-664-68
C 9	98.6	10.6	641	17	US-10-333-184-66
C 10	98.6	10.6	996	9	US-09-770-445-241
C 11	98.2	10.5	524	18	US-10-021-323-12510

Sequence 15571, A	US-10-021-323-16571	18	544	10.5	98.2	12	C 13
Sequence 6450, Ap	US-10-767-795-6450	18	619	10.5	98.2	13	C 14
Sequence 12991, A	US-10-425-114-12991	17	765	10.4	97	14	C 15
Sequence 172656, A	US-10-425-115-172656	18	1822	10.4	97	15	C 16
Sequence 11357, A	US-10-021-323-11357	18	514	10.0	93.4	16	C 17
Sequence 142308, A	US-10-424-599-142308	17	1954	10.0	93.4	17	C 18
Sequence 14386, A	US-10-424-599-14386	17	1791	9.9	92.6	18	C 19
Sequence 15101, A	US-10-021-323-15101	18	443	9.7	90.6	19	C 20
Sequence 16704, A	US-10-021-323-16704	18	444	9.7	90.6	20	C 21
Sequence 16920, A	US-10-021-323-16920	18	512	9.7	90.6	21	C 22
Sequence 10302, A	US-10-021-323-10302	18	528	9.7	90.6	22	C 23
Sequence 12560, A	US-10-021-323-12560	18	531	9.7	90.6	23	C 24
Sequence 6892, Ap	US-10-021-323-6892	18	569	9.7	90.6	24	C 25
Sequence 10831, A	US-10-021-323-10831	18	573	9.7	90.6	25	C 26
Sequence 13101, A	US-10-021-323-13101	18	574	9.7	90.6	26	C 27
Sequence 16704, A	US-10-021-323-16704	18	597	9.7	90.6	27	C 28
Sequence 15060, A	US-10-424-599-50160	18	608	9.7	90.6	28	C 29
Sequence 8545, Ap	US-10-021-323-8545	18	1762	9.7	90.6	29	C 30
Sequence 6451, Ap	US-10-767-795-6451	18	530	9.6	89.6	30	C 31
Sequence 17161, A	US-10-021-323-17161	18	530	9.6	89.6	31	C 32
Sequence 10449, A	US-10-021-323-10449	18	559	9.6	89.6	32	C 33
Sequence 17838, A	US-10-021-323-17838	18	582	9.6	89.6	33	C 34
Sequence 5579, Ap	US-10-021-323-5579	18	592	9.6	89.6	34	C 35
Sequence 1845, Ap	US-09-938-842A-1845	9	1350	9.5	89	35	C 36
Sequence 134384, A	US-09-938-842A-1845	11	1350	9.5	89	36	C 37
Sequence 142276, A	US-10-424-599-134384	17	1929	9.5	88.6	37	C 38
Sequence 142286, A	US-10-424-599-142276	17	621	9.5	88.6	38	C 39
Sequence 10860, A	US-10-424-599-142286	17	1840	9.5	88.6	39	C 40
Sequence 79417, A	US-10-425-114-10860	17	1705	9.5	88.4	40	C 41
Sequence 13438, A	US-10-424-599-79417	17	1815	9.5	88.4	41	C 42
Sequence 135, App	US-10-021-323-13438	18	494	9.4	87.8	42	C 43
Sequence 134566, A	US-10-133-013-135	16	588	9.4	87.8	43	C 44
Sequence 175282, A	US-10-424-599-134566	17	890	9.4	87.8	44	C 45
Sequence 29586, A	US-10-425-115-175282	17	968	9.4	87.8	45	C 46
Sequence 3923, Ap	US-10-425-114-29586	17	1704	9.4	87.8	46	C 47
Sequence 169190, A	US-10-260-238-3923	18	1707	9.4	87.8	47	C 48
Sequence 175283, A	US-10-425-115-169190	18	1845	9.4	87.8	48	C 49
Sequence 93240, A	US-10-425-115-175283	18	2066	9.4	87.8	49	C 50
Sequence 169187, A	US-10-437-963-93240	18	2072	9.4	87.8	50	C 51
Sequence 169188, A	US-10-425-115-169187	18	2072	9.4	87.8	51	C 52
Sequence 59521, A	US-10-425-115-169188	17	2075	9.4	87.8	52	C 53
Sequence 35442, A	US-10-424-599-58521	17	2311	9.4	87.4	53	C 54
Sequence 35804, A	US-10-425-114-35442	17	761	9.4	87.4	54	C 55
Sequence 133716, A	US-10-425-114-35804	17	900	9.4	87.4	55	C 56
Sequence 50161, A	US-10-424-599-133716	17	1853	9.4	87.4	56	C 57
Sequence 140251, A	US-10-424-599-50161	17	1907	9.4	87.4	57	C 58
Sequence 133718, A	US-10-425-115-140251	18	2030	9.4	87.4	58	C 59
Sequence 167, App	US-10-424-599-133718	17	2094	9.4	87.4	59	C 60
Sequence 2342, Ap	US-09-770-961-167	10	513	9.3	87	60	C 61
Sequence 58517, A	US-09-938-842A-2342	9	1344	9.3	87	61	C 62
Sequence 4982, Ap	US-09-938-842A-2342	11	1344	9.3	87	62	C 63
Sequence 427, App	US-10-424-599-58517	17	1936	9.3	86.4	63	C 64
Sequence 30272, A	US-10-260-238-4982	17	623	9.3	86.4	64	C 65
Sequence 3621, Ap	US-10-333-184-427	17	405	9.2	86.2	65	C 66
Sequence 171779, A	US-10-425-114-30272	17	1142	9.2	86.2	66	C 67
Sequence 231, App	US-10-425-114-3621	17	1646	9.2	86.2	67	C 68
Sequence 30543, A	US-10-425-115-171779	18	1735	9.2	86.2	68	C 69
Sequence 31497, A	US-09-770-423-231	9	426	9.2	86	69	C 70
Sequence 140247, A	US-10-369-493-30543	17	567	9.2	85.8	70	C 71
Sequence 140239, A	US-10-425-114-31497	17	834	9.2	85.8	71	C 72
Sequence 49528, A	US-10-425-115-140247	18	1081	9.2	85.8	72	C 73
Sequence 140245, A	US-10-425-115-140239	17	1334	9.2	85.8	73	C 74
Sequence 140242, A	US-10-437-963-49528	18	1670	9.2	85.8	74	C 75
Sequence 140250, A	US-10-437-963-49528	18	1837	9.2	85.8	75	C 76
Sequence 3828, Ap	US-10-425-115-140242	18	1855	9.2	85.8	76	C 77
Sequence 429, App	US-10-425-115-140242	18	1915	9.2	85.8	77	C 78
Sequence 5066, Ap	US-10-425-115-140250	18	2115	9.2	85.8	78	C 79
Sequence 41161, A	US-10-260-238-3828	17	540	9.1	84.8	79	C 80
Sequence 12013, A	US-10-333-184-429	17	852	9.1	84.8	80	C 81
Sequence 3826, Ap	US-10-739-930-5066	18	1762	9.1	84.8	81	C 82
Sequence 3607, Ap	US-10-437-963-41161	18	1930	9.1	84.8	82	C 83
	US-10-021-323-12013	18	480	9.1	84.6	83	C 84
	US-10-260-238-3607	17	666	9.1	84.6	84	C 85
	US-10-260-238-3607	17	688	9.1	84.6	85	C 86

Sequence 3824, Ap
 Sequence 3618, Ap
 Sequence 3619, Ap
 Sequence 5056, Ap
 Sequence 5060, Ap
 Sequence 5061, Ap
 Sequence 13177, A
 Sequence 134568,
 Sequence 7739, Ap
 Sequence 3829, Ap
 Sequence 3827, Ap
 Sequence 5063, Ap
 Sequence 170926,
 Sequence 16911, A
 Sequence 6957, Ap
 Sequence 6884, Ap

ALIGNMENTS

RESULT 1
 US-09-938-842A-622/c
 ; Sequence 622, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPT300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 622
 ; LENGTH: 1335
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-622

Query Match 11.2%; Score 104; DB 9; Length 1335;
 Best Local Similarity 88.3%; Pred. No. 8.8e-13;
 Matches 113; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 4 CTGTAGCATCTTGATATTGCTGATCTCAGCCACAGATCGTTTCATGTTACTCTCTGCTT 63
 Db 1288 CTGTAGCATCTTGATATTGCTGATCTCAGCCACAGATCATTCATGTTACTCTCTGCTT 1229
 QY 64 CATTAAACTCCATCTCGTCATTCCTCTTCTGTGTACCAATGCAAGAAAGCTTATCTCA 123
 Db 1228 CAGTGAACCTCCATCTCGTCCATTCCTCTCTGTGTACCAATGCAAGAAAGCTTCTCC 1169
 QY 124 ACATCAGG 131
 Db 1168 TGAACATG 1161

RESULT 2
 US-09-938-842A-622/c
 ; Sequence 622, Application US/09938842A
 ; Publication No. US20040009476A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPT300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 622
 ; LENGTH: 1335
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-622

Query Match 11.2%; Score 104; DB 11; Length 1335;
 Best Local Similarity 88.3%; Pred. No. 8.8e-13;
 Matches 113; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 4 CTGTAGCATCTTGATATTGCTGATCTCAGCCACAGATCGTTTCATGTTACTCTCTGCTT 63
 Db 1288 CTGTAGCATCTTGATATTGCTGATCTCAGCCACAGATCATTCATGTTACTCTCTGCTT 1229
 QY 64 CATTAAACTCCATCTCGTCATTCCTCTTCTGTGTACCAATGCAAGAAAGCTTATCTCA 123
 Db 1228 CAGTGAACCTCCATCTCGTCCATTCCTCTCTGTGTACCAATGCAAGAAAGCTTCTCC 1169
 QY 124 ACATCAGG 131
 Db 1168 TGAACATG 1161

RESULT 3
 US-09-770-445-8
 ; Sequence 8, Application US/09770445
 ; Patent No. US20020023281A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gorlach, Jorn
 ; APPLICANT: An, Yong-Qiang
 ; APPLICANT: Hamilton, Carol M.
 ; APPLICANT: Price, Jennifer L.
 ; APPLICANT: Raines, Tracy M.
 ; APPLICANT: Yu, Yang
 ; APPLICANT: Rameaka, Joshua G.
 ; APPLICANT: Page, Amy
 ; APPLICANT: Matthew, Abraham V.
 ; APPLICANT: Ledford, Brooke L.
 ; APPLICANT: Woessner, Jeffrey P.
 ; APPLICANT: Haas, William David
 ; APPLICANT: Garcia, Carlos A.
 ; APPLICANT: Krickler, Maja
 ; APPLICANT: Slader, Ted
 ; APPLICANT: Davis, Keith R.
 ; APPLICANT: Allen, Keith
 ; APPLICANT: Hoffman, Neil
 ; APPLICANT: Hurlban, Patrick
 ; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
 ; FILE REFERENCE: 2023US (PARA-012PRV)
 ; CURRENT APPLICATION NUMBER: US/09/770,445
 ; CURRENT FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: US 60/178,472
 ; PRIOR FILING DATE: 2000-01-27
 ; NUMBER OF SEQ ID NOS: 999
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 1469
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:25:44 ; Search time 614 Seconds
(without alignment)
8985.666 Million cell updates/sec

Title: US-10-009-340-1
Perfect score: 932
Sequence: 1 cagctgttagctctgtgat.....agagcaacaacctggcacag 932

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	932	100.0	932	5	AAT28949 Arabidops
2	104	11.2	1335	6	ABZ12817 Arabidops
3	104	11.2	1469	6	ABN98240 Arabidops
4	104	11.2	1632	3	AAC47401 Arabidops
5	104	11.2	1636	3	AAC37579 Arabidops
6	100.8	10.8	772	6	ABK31017 Plant dwa
7	99.2	10.6	469	6	ABK30849 Plant dwa
8	98.6	10.6	445	6	ABL94041 Arabidops
9	98.6	10.6	586	6	ABK30889 Arabidops
10	98.6	10.6	641	6	ABK30652 Plant dwa
11	98.6	10.6	641	6	ABK30652 Plant dwa
12	98.6	10.6	996	6	ABN98473 Arabidops
13	98.6	10.6	1601	3	AAC40829 Arabidops
14	98.6	10.6	1604	3	AAC47026 Arabidops
15	98.6	10.6	1643	3	AAC48904 Arabidops
16	98.2	10.5	524	13	ACN57729 Cotton gy
17	98.2	10.5	544	13	ACN61790 Cotton gy
18	98.2	10.5	619	13	ADR56669 Cotton cd
19	97	10.4	882	3	AAC48727 Arabidops
20	97	10.4	1353	12	ADN73732 Thale cre

c 21	97	10.4	1437	3	AAC46796 Arabidops
c 22	97	10.4	1621	3	AAC47378 Arabidops
c 23	97	10.4	1624	3	AAC35010 Arabidops
c 24	97	10.4	1706	3	AAC37105 Arabidops
c 25	95.4	10.2	1430	3	AAC46947 Arabidops
c 26	94	10.1	150	3	AAC49119 Arabidops
c 27	93.8	10.1	1353	3	AAC49774 Arabidops
c 28	93.8	10.1	1611	3	AAC49782 Arabidops
c 29	93.8	10.1	1643	3	AAC49781 Arabidops
c 30	93.8	10.1	1649	3	AAC40306 Arabidops
c 31	93.4	10.0	514	13	ACN56576 Cotton gy
c 32	90.6	9.7	443	13	ACN60320 Cotton gy
c 33	90.6	9.7	444	13	ACN61923 Cotton gy
c 34	90.6	9.7	512	13	ACN62139 Cotton gy
c 35	90.6	9.7	528	13	ACN55521 Cotton gy
c 36	90.6	9.7	531	13	ACN57779 Cotton gy
c 37	90.6	9.7	569	13	ACN52111 Cotton an
c 38	90.6	9.7	573	13	ACN56050 Cotton an
c 39	90.6	9.7	574	13	ACN56613 Cotton an
c 40	90.6	9.7	608	13	ACN53764 Cotton an
c 41	90.6	9.7	1762	13	ADR56670 Cotton cd
c 42	89.6	9.6	530	13	ACN62380 Cotton gy
c 43	89.6	9.6	559	13	ACN55668 Cotton an
c 44	89.6	9.6	582	13	ACN63057 Cotton ca
c 45	89.6	9.6	592	13	ACN50798 Cotton an
c 46	89	9.5	1350	6	ABZ14040 Arabidops
c 47	89	9.5	1496	3	AAC48427 Arabidops
c 48	89	9.5	1601	3	AAC34041 Arabidops
c 49	87.8	9.4	414	3	AAC46724 Arabidops
c 50	87.8	9.4	476	3	AAC46728 Arabidops
c 51	87.8	9.4	494	13	ACN58657 Arabidops
c 52	87.8	9.4	588	10	ADI02654 Human cdn
c 53	87.8	9.4	1707	12	ADJ42823 Plant cdn
c 54	87	9.3	441	3	AAC41663 Arabidops
c 55	87	9.3	513	10	ADE81396 Arabidops
c 56	87	9.3	1344	6	ABZ14537 Arabidops
c 57	87	9.3	1564	3	AAC47973 Arabidops
c 58	86.4	9.3	475	3	AAC35886 Arabidops
c 59	86.4	9.3	623	12	ADJ43982 Plant cdn
c 60	86.2	9.2	405	6	ABK31013 Plant dwa
c 61	86	9.2	426	8	ABK62116 Arabidops
c 62	86	9.2	2833	13	ADS91274 beta-1 tu
c 63	85.8	9.2	567	13	ADS54869 Bacterial
c 64	85	9.1	1347	2	AAT41699 Onchocerc
c 65	84.8	9.1	540	12	ADJ42828 Plant cdn
c 66	84.8	9.1	852	6	ABK31015 Plant dwa
c 67	84.8	9.1	1623	6	ABK15698 Cotton cd
c 68	84.6	9.1	480	13	ACN57232 Cotton gy
c 69	84.6	9.1	585	13	ADQ50711 Novel can
c 70	84.6	9.1	666	12	ADJ42826 Plant cdn
c 71	84.6	9.1	688	12	ADJ42607 Plant cdn
c 72	84.6	9.1	868	12	ADJ42824 Plant cdn
c 73	84.6	9.1	884	12	ADJ42618 Plant cdn
c 74	84.6	9.1	916	12	ADJ42619 Plant cdn
c 75	83.6	9.0	533	9	ACL18164 DNA clone
c 76	83.4	8.9	1347	2	AAT41698 Arabidops
c 77	83	8.9	488	9	ACL18168 DNA clone
c 78	83	8.9	540	9	ACL18160 DNA clone
c 79	83	8.9	582	9	ACL18144 DNA clone
c 80	83	8.9	598	13	ACN52958 Cotton an
c 81	83	8.9	601	9	ACL18165 DNA clone
c 82	83	8.9	609	9	ACL18177 DNA clone
c 83	83	8.9	617	9	ACL18179 DNA clone
c 84	83	8.9	620	9	ACL18173 DNA clone
c 85	83	8.9	622	9	ACL18178 DNA clone
c 86	83	8.9	663	9	ACL18182 DNA clone
c 87	83	8.9	668	9	ACL18181 DNA clone
c 88	83	8.9	669	9	ACL18170 DNA clone
c 89	83	8.9	680	9	ACL18174 DNA clone
c 90	83	8.9	683	9	ACL18169 DNA clone
c 91	83	8.9	693	9	ACL18163 DNA clone
c 92	83	8.9	693	9	ACL18161 DNA clone
c 93	83	8.9	701	9	ACL18180 DNA clone

94 83 8.9 712 9 ACL18171 Acl18171 DNA clone
 95 83 8.9 720 9 ACL18176 Acl18176 DNA clone
 c 96 83 8.9 736 12 ADJ42829 Adj42829 Plant CDN
 c 97 83 8.9 815 12 ADJ42827 Adj42827 Plant CDN
 98 82.6 8.9 466 13 ACN62130 ACN62130 Cotton gy
 99 82.6 8.9 585 13 ACN52176 ACN52176 Cotton an
 100 82.6 8.9 596 13 ACN52103 ACN52103 Cotton an

ALIGNMENTS

RESULT 1

AAF28949 ID AAF28949 standard; DNA; 932 BP.
 XX AC
 XX AAF28949;
 DT 02-JUL-2002 (first entry)
 XX Arabidopsis thaliana fatty acid hydroxylase gene promoter.
 DE Arabidopsis thaliana
 KW promoter; gene expression; plant; plant tissue; seed; FAH; frost;
 KW fatty acid hydroxylase; stress resistance; pest; pathogen; heat; drought;
 KW metabolism; antisense; transgene; ds.
 XX Arabidopsis thaliana.
 OS WO200077223-A1.
 PN 21-DEC-2000.
 PD 08-JUN-2000; 2000WO-FR001574.
 PF 10-JUN-1999; 99FR-00007362.
 PR (INRG) INST NAT RECH AGRONOMICQUE.
 XX Dubreucq B, Lepiniec L, Caboche M;
 FI WPI; 2001-080694/09.
 DR New plant promoter from the fatty acid hydroxylase gene, useful for
 PT expressing transgenes that e.g. improve stress resistance, in all tissues
 PT except ripening and dry seed.
 XX Claim 3; Page 25; 33pp; French.

XX The invention relates to the isolation of a novel promoter that provides
 CC expression of a gene in all plant tissues other than ripening and dry
 CC seeds is at least 80% identical with at least part of the Arabidopsis FAH
 CC (fatty acid hydroxylase) gene promoter. The promoter and related
 CC promoters from other plants, are used to direct expression of the gene
 CC that (i) improve resistance to stress (e.g. pests, pathogens, heat, frost,
 CC drought etc.); (ii) to modulate development; (iii) encode enzymes
 CC involved in metabolic processes or (iv) inhibit a target gene (by co-
 CC suppression or antisense techniques). Transgenes are expressed in all
 CC tissues of the adult plant but not its expression product is not present
 CC in seeds (so such seeds are likely to be more acceptable to consumers).
 CC This sequence represents the promoter region of the FAH gene from A.
 CC thaliana

SQ Sequence 932 BP; 318 A; 174 C; 141 G; 299 T; 0 U; 0 Other;
 Query Match 100.0%; Score 932; DB 5; Length 932;
 Best Local Similarity 100.0%; Pred. No. 4.5e-191;
 Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAGCTGTAGCATCTTGATATTCGTGATCTAGTACGACCAAGATCGTTTCATGTTACTCTCTG 60
 DB 1 CAGCTGTAGCATCTTGATATTCGTGATCTAGTACGACCAAGATCGTTTCATGTTACTCTCTG 60
 QY 61 CTTCAATAACTCCATCTCGTCCATCTCTCTCTGTGTACCAATGTCAGAAAGCTTATC 120

DB 61 CTTCAATAACTCCATCTCGTCCATCTCTCTCTGTGTACCAATGTCAGAAAGCTTATC 120
 QY 121 TCAACATCAGCGCTGATATACCAATATCTTACTTCTTTTACATTTTGGAAATGGAACCAA 180
 DB 121 TCAACATCAGCGCTGATATACCAATATCTTACTTCTTTTACATTTTGGAAATGGAACCAA 180
 QY 181 CCCATTTTCTCGAAAAAGTGCTAAACCAACATTTGATTAAACCGTATCCTACTACTTTTC 240
 DB 181 CCCATTTTCTCGAAAAAGTGCTAAACCAACATTTGATTAAACCGTATCCTACTACTTTTC 240
 QY 241 ATTTCTATCTTCTGTTTCATTTATGCTGACTATTTAAAGTCCGTTGTCCTTAAGTT 300
 DB 241 ATTTCTATCTTCTGTTTCATTTATGCTGACTATTTAAAGTCCGTTGTCCTTAAGTT 300
 QY 301 AGACATAAAGACAAAGACTAAATCAATTTGTCATCACACCGCGTCGTCGAGTGAGCTATA 360
 DB 301 AGACATAAAGACAAAGACTAAATCAATTTGTCATCACACCGCGTCGTCGAGTGAGCTATA 360
 QY 361 TTAATCGTGGATTTTAAAGCAATTAAGAAACATTTCTATAGTACTAAAGCAATAAATAAT 420
 DB 361 TTAATCGTGGATTTTAAAGCAATTAAGAAACATTTCTATAGTACTAAAGCAATAAATAAT 420
 QY 421 TATAATCAAAACACTATCTTGCACACTCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 421 TATAATCAAAACACTATCTTGCACACTCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 ATAAAGCGCGCATCAAAATCTTAAAGCAATTAAGCAATTAAGCAATTAAGCAATTAAGCAAT 540
 DB 481 ATAAAGCGCGCATCAAAATCTTAAAGCAATTAAGCAATTAAGCAATTAAGCAATTAAGCAAT 540
 QY 541 ATTTTACTCAATAAGAAATCGAAAGTATGATTTATTTATCTAGCTGCCACAACTCTCGAAT 600
 DB 541 ATTTTACTCAATAAGAAATCGAAAGTATGATTTATTTATCTAGCTGCCACAACTCTCGAAT 600
 QY 601 TTAATATTTTACTCAAGAAAGACCGACTTTTAAATCTTGAATTTTCTTCTTCTTCTTCTTCT 660
 DB 601 TTAATATTTTACTCAAGAAAGACCGACTTTTAAATCTTGAATTTTCTTCTTCTTCTTCTTCT 660
 QY 661 AATGATTAAGCACTCAATAAATCTTTCACATTTTGGCAGAGCAATTAATTCGAA 720
 DB 661 AATGATTAAGCACTCAATAAATCTTTCACATTTTGGCAGAGCAATTAATTCGAA 720
 QY 721 GTCTAAAGTAAATCGTCCACACAGTGTAGTATCTAGTATTTTCTTCTTCTTCTTCTTCT 780
 DB 721 GTCTAAAGTAAATCGTCCACACAGTGTAGTATCTAGTATTTTCTTCTTCTTCTTCTTCT 780
 QY 781 ATAAAGTTGAATTTGTAAT 840
 DB 781 ATAAAGTTGAATTTGTAAT 840
 QY 841 TGGGAATCCTATATAGTGGCAGCAGCAAGTGAACAAGCAGGCTATGCAACAGCTTAATGTATG 900
 DB 841 TGGGAATCCTATATAGTGGCAGCAGCAAGTGAACAAGCAGGCTATGCAACAGCTTAATGTATG 900
 QY 901 AAGAGAGCCAAAGAGCAACAACTGGCACAG 932
 DB 901 AAGAGAGCCAAAGAGCAACAACTGGCACAG 932

RESULT 2

ABZ12817/c ID ABZ12817 standard; DNA; 1335 BP.
 XX AC
 XX ABZ12817;
 XX 21-JAN-2003 (first entry)
 XX Arabidopsis thaliana stress regulated gene SEQ ID NO 622.
 DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 KW Arabidopsis thaliana.
 OS Arabidopsis thaliana.

Bausch, S.
101609340
Seq. 151

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:48:24 ; Search time 4373 seconds
(without alignments)
10327.066 Million cell updates/sec

Title: US-10-009-340-1
Perfect score: 932
Sequence: 1 cagctgtacatcttgatat.....agagcaacaacctggcacag 932

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	932	100.0	932	6	AX058259 Sequence
2	932	100.0	69817	8	AC003096 Arabidops
3	105.6	11.3	1334	8	BT012803 Lycopersi
4	104	11.2	1335	6	AX050927 Sequence
5	104	11.2	1563	8	AY128337 Arabidops
6	104	11.2	1636	8	AY087594 Arabidops
7	104	11.2	2247	8	ATHTUB9B Arabidops
8	104	11.2	91570	8	ATHTUB9B Arabidops
9	104	11.2	197568	8	ATHTUB9B Arabidops
10	100.8	10.8	773	6	AX364424 Sequence
11	99.2	10.6	469	6	AX364256 Sequence
12	98.6	10.6	586	6	AX364296 Sequence
13	98.6	10.6	641	6	AX364059 Sequence
14	98.6	10.6	641	6	AX366149 Sequence
15	98.6	10.6	1366	8	AY059075 Arabidops
16	98.6	10.6	1731	8	AY035141 Arabidops
17	98.6	10.6	2936	8	ATHBT4A Arabidops
18	98.6	10.6	60583	8	AB011475 Arabidops
19	97.6	10.5	120562	8	AC124958 Medicago

C 20	97.6	10.5	140922	2	AC144474
C 21	97.2	10.4	412	11	BV165737
C 22	97	10.4	1344	8	BT005372
C 23	97	10.4	1350	6	BT001121
C 24	97	10.4	1353	6	CQ805216
C 25	97	10.4	1395	8	AF370568
C 26	97	10.4	1538	8	AF370568
C 27	97	10.4	1608	8	AK117262
C 28	97	10.4	1624	8	AY085710
C 29	97	10.4	1652	8	AY060550
C 30	97	10.4	1737	8	AF367332
C 31	97	10.4	2374	8	ATHTUB7B
C 32	97	10.4	2764	8	ATHTUB3B
C 33	97	10.4	2951	8	ATU35049
C 34	97	10.4	3094	8	ATHTUBB
C 35	97	10.4	18419	8	AF049870
C 36	97	10.4	78145	8	AB009053
C 37	97	10.4	95870	8	AC007396
C 38	97	10.4	129547	8	AC006434
C 39	97	10.4	131743	8	AC004561
C 40	96.4	10.3	413	11	BV165738
C 41	96	10.3	876	8	APTUBB3A
C 42	95.6	10.3	139440	2	AC137926
C 43	95.4	10.2	1281	8	MSA319667
C 44	95.4	10.2	1350	8	AY040074
C 45	95.4	10.2	1578	8	BT014148
C 46	95.4	10.2	1649	8	AK117925
C 47	95.4	10.2	1661	8	AF360260
C 48	95.4	10.2	2061	8	ATHTUB6B
C 49	95.4	10.2	79590	8	AB007727
C 50	95.4	10.2	98506	8	ATT2L20
C 51	94	10.1	1335	3	S72098
C 52	93.8	10.1	1441	8	AY128785
C 53	93.8	10.1	1442	8	BT002587
C 54	93.8	10.1	1596	8	BT013153
C 55	93.8	10.1	1645	8	AY087859
C 56	93.8	10.1	1647	8	AY062614
C 57	93.8	10.1	1656	8	AY056411
C 58	93.8	10.1	1886	8	AY062742
C 59	93.8	10.1	2406	8	ATHTUB2B
C 60	93.8	10.1	2699	8	ATHTUB2B
C 61	93.8	10.1	55151	8	AB020751
C 62	93.4	10.0	409	8	D78496
C 63	93.4	10.0	2100	8	LAU47660
C 64	92.8	10.0	1602	8	D63137
C 65	92.6	9.9	1677	8	AF059287
C 66	90.8	9.7	417	8	NTU91564
C 67	90.6	9.7	389	11	BV165741
C 68	90.6	9.7	398	11	BV165739
C 69	90.6	9.7	402	11	BV165742
C 70	90.6	9.7	1511	8	PSGTUB2
C 71	90.6	9.7	1789	8	AY704406
C 72	90.6	9.7	48314	8	AP004982
C 73	90.2	9.7	1633	8	BT013065
C 74	90.2	9.7	3023	8	PSGTUB1
C 75	89.8	9.6	2445	5	XLU15444
C 76	89.8	9.6	2460	5	BC049004
C 77	89.8	9.6	2492	5	BC074549
C 78	89.6	9.6	1524	3	EOBETA
C 79	89.6	9.6	1634	8	CABTUB
C 80	89	9.5	293	8	AJ718573
C 81	89	9.5	293	8	AJ718573
C 82	89	9.5	405	11	BV165740
C 83	89	9.5	1350	6	AX0507150
C 84	89	9.5	1422	8	D63138
C 85	89	9.5	1481	8	AY081473
C 86	89	9.5	1583	8	BT013893
C 87	89	9.5	1648	8	AY054893
C 88	89	9.5	2187	8	ATHTUB8B
C 89	89	9.5	82415	8	AB005244
C 90	88.6	9.5	1524	8	STBETUB2
C 91	88.6	9.5	1557	8	STBETUB1
C 92	88	9.4	1413	8	PSGTUB3

AC144474	Medicago
BV165737	TBB2-1 PC
BT005372	Arabidops
BT001121	Arabidops
CQ805216	Sequence
AF370568	Arabidops
AF370568	Arabidops
AK117262	Arabidops
AY085710	Arabidops
AY060550	Arabidops
AF367332	Arabidops
M84704	Arabidops
M84701	Arabidops
U35049	Arabidops
M20405	Arabidops
AF049870	Arabidops
AB009053	Arabidops
AC007396	Genomic s
AC006434	Genomic s
AC004561	Arabidops
BV165738	TBB2-2 PC
X69187	A.pynllitid
AC137926	Brassica
AJ319667	Medicago
AY040074	Arabidops
BT014148	Lycopersi
AK117925	Arabidops
AF360260	Arabidops
M84703	Arabidops
AB007727	Arabidops
AL592312	Arabidops
S72098	Euplotes fo
AY128785	Arabidops
BT002587	Arabidops
BT013153	Lycopersi
AY087859	Arabidops
AY062614	Arabidops
AY056411	Arabidops
AY062742	Arabidops
AK117431	Arabidops
M84700	Arabidops
AB020751	Arabidops
D78496	Brassica ra
U47660	Lupinus alb
D63137	Zinnia eleg
AF059287	Eleusine
U91564	Nicotiana t
BV165741	TBB2-14 PC
BV165739	TBB2-5 PC
BV165742	TBB2-15 P
X54845	Pea gTUB2 m
AY704406	Quercus p
AP004982	Lotus cor
BT013065	Lycopersi
X54844	P.sativum g
U15444	Xenopus lae
BC049004	Xenopus l
BC074549	Xenopus t
BC074549	Xenopus t
EOBETA	E.octocarin
X98406	C.arrietinum
AJ718573	Nicotiana
AJ718574	Nicotiana
BV165740	TBB2-6 PC
AX0507150	Sequence
D63138	Zinnia eleg
AY081473	Arabidops
BT013893	Lycopersi
AY054893	Arabidops
M84705	Arabidops
AB005244	Arabidops
Z33402	S.tuberosum
Z3382	S.tuberosum
X54846	P.sativum g

c 93 88 9.4 2424 8 LALB1
 c 94 87.8 9.4 1533 8 APTUBBIA
 c 95 87.8 9.4 1583 8 MZBETUB7A
 c 96 87.8 9.4 1584 8 BT013141
 c 97 87.8 9.4 1695 8 AY382291
 c 98 87.8 9.4 1741 8 AK120180
 c 99 87.8 9.4 1854 8 RICBETAT
 c 100 87.8 9.4 178158 8 AC084320

ALIGNMENTS

RESULT 1
 LOCUS AX058259 AX058259 932 bp DNA linear PAT 17-JAN-2001
 DEFINITION Sequence 1 from Patent WO0077223.
 ACCESSION AX058259
 VERSION AX058259.1 GI:12310760

SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1
 Dubreucq, B., Lepiniec, L. and Caboche, M.
 Promoter enabling transgene expression in the whole plant except in
 the seed

JOURNAL Patent: WO 0077223-A 1 21-DEC-2000;
 INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA)

FEATURES
 source Location/Qualifiers

1..932
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"
 /notes="promoteur de la FAH chez Arabidopsis thaliana."

ORIGIN

Query Match 100.0%; Spore 932; DB 6; Length 932;
 Best Local Similarity 100.0%; Ered. No. 1e-167;
 Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCTGACATCTGATATGCTGATCTCAGCCACCAAGATCGTTTCATGTTACTCTCTG 60
 DB 1 CAGCTGATGATCTGATATGCTGATCTCAGCCACCAAGATCGTTTCATGTTACTCTCTG 60
 QY 61 CTTCAATTAACCTCCATCTGTCATCTCTCTCTGTCACCAATGCAAGATGCTTATC 120
 DB 61 CTTCAATTAACCTCCATCTGTCATCTCTCTCTGTCACCAATGCAAGATGCTTATC 120
 QY 121 TCAACATCAGGCTGATATACCAATATCTTCTCTCTGTCACCAATGCAAGATGCTTATC 180
 DB 121 TCAACATCAGGCTGATATACCAATATCTTCTCTCTGTCACCAATGCAAGATGCTTATC 180
 QY 181 CCCATTTTCTGGAAGAGTCTAATCAATCTTCTCTCTGTCACCAATGCAAGATGCTTATC 240
 DB 181 CCCATTTTCTGGAAGAGTCTAATCAATCTTCTCTCTGTCACCAATGCAAGATGCTTATC 240
 QY 241 ATTTCTATCTCTGTTTCATATGCTGATATTAAGCTCCGTTGTCATATCTTAAAGTT 300
 DB 241 ATTTCTATCTCTGTTTCATATGCTGATATTAAGCTCCGTTGTCATATCTTAAAGTT 300
 QY 301 AGACATAAAGACCAAGACTAATCAATGTCATCAACGCTCGTCGAGTACGCTATA 360
 DB 301 AGACATAAAGACCAAGACTAATCAATGTCATCAACGCTCGTCGAGTACGCTATA 360
 QY 361 TTAATCGTGATTTTAAAGCAATTAAGCAATCTTCTATAGTACTAAAGCAATTAATAT 420
 DB 361 TTAATCGTGATTTTAAAGCAATTAAGCAATCTTCTATAGTACTAAAGCAATTAATAT 420
 QY 421 TATAATCAACACTATGCTTGACACTGGTCAACGCTGATCTGGTAGTGAATGTTCTACATC 480

Db 421 TATAATCAACACTATGCTTGACACTGGTCAACGCTGATCTGGTAGTGAATGTTCTACATC 480
 QY 481 ATAAGAGGCGCGATCAAAATCT 540
 Db 481 ATAAGAGGCGCGATCAAAATCT 540
 QY 541 ATTTTACTCAATTAAGCAATCGAAAGTATGATTTATCTAGTCTGCAATCTTCCGAAT 600
 Db 541 ATTTTACTCAATTAAGCAATCGAAAGTATGATTTATCTAGTCTGCAATCTTCCGAAT 600
 QY 601 TTAATATTTTACTCAAGAGAGACCGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
 Db 601 TTAATATTTTACTCAAGAGAGACCGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
 QY 661 AATGATTTAAAGCAGTCAATATAATCTTTTGACATTTGTCGCAAGACCAATAATTCGAA 720
 Db 661 AATGATTTAAAGCAGTCAATATAATCTTTTGACATTTGTCGCAAGACCAATAATTCGAA 720
 QY 721 GTCTAAATGTAATCGTCCACACAGTGTATGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 Db 721 GTCTAAATGTAATCGTCCACACAGTGTATGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 QY 781 AATGATTTAAAGCAGTCAATATAATCTTTTGACATTTGTCGCAAGACCAATAATTCGAA 840
 Db 781 AATGATTTAAAGCAGTCAATATAATCTTTTGACATTTGTCGCAAGACCAATAATTCGAA 840
 QY 841 TGGGAATCTCTATAAGTGGCAGCAGCAAGTGCACCAAGTGCACCAAGTGCACCAAGTGCACCA 900
 Db 841 TGGGAATCTCTATAAGTGGCAGCAGCAAGTGCACCAAGTGCACCAAGTGCACCAAGTGCACCA 900
 QY 901 AAGAGAGCCCAAGAGCAACACCTGGGCACAG 932
 Db 901 AAGAGAGCCCAAGAGCAACACCTGGGCACAG 932

RESULT 2

AC003096/c
 LOCUS Arabidopsis thaliana chromosome 2 clone T29F13 map ve016, complete
 DEFINITION sequence.
 ACCESSION AC003096
 VERSION AC003096.3 GI:20197018
 KEYWORDS HTG.

SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 69817)
 Rounsley, S.D., Lin, X., Ketchum, K.A., Crosby, M.L., Brandon, R.C.,
 Sykes, S.M., Kaul, S., Mason, T.M., Kerlavage, A.R., Adams, M.D.,
 Somerville, C.R. and Venter, J.C.

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 69817)
 Lin, X.

DIRECT SUBMISSION
 TITLE Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
 JOURNAL Medical Center Dr., Rockville, MD 20850, USA

REFERENCE 3 (bases 1 to 69817)
 Town, C.D. and Kaul, S.

DIRECT SUBMISSION
 TITLE Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
 JOURNAL Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org

COMMENT On Apr 18, 2002 this sequence version replaced gi:598389.
 FEATURES Location/Qualifiers

source
 1..69817
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="2"
 /map="ve016"
 /clone="T29F13"

Query Match	100.0%;	Score 932;	DB 8;	Length 69817;
Best Local Similarity	100.0%;	Pred. No. 4.8e-160;		
Matches 932;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CAGCTGTAGCATCTTGATATTGCTGATACCTACGACCACAGATCGTTCAITGTTACTCTCTG	60	
Db	7798	CAGCTGTAGCATCTTGATATTGCTGATACCTACGACCACAGATCGTTCAITGTTACTCTCTG	7739	
Qy	61	CTTCATTAAACTCCACTCTCGTCCATTCTTCTCTGTGTACCAATGCAAGAAGCTTATC	120	
Db	7738	CTTCATTAAACTCCACTCTCGTCCATTCTTCTCTGTGTACCAATGCAAGAAGCTTATC	7679	
Qy	121	TCAACATCAGGCTGATATACCAATATCTTATCTTTTACATTTGTGAAATGGAAACAA	180	
Db	7678	TCAACATCAGGCTGATATACCAATATCTTACTTCTTTTACATTTGTGAAATGGAAACAA	7619	
Qy	181	CCCATTTTTCGAAAAGTGCTTAACCAACATTTGATTTAAACGGTATCACTACTACTTTC	240	
Db	7618	CCCATTTTTCGAAAAGTGCTTAACCAACATTTGATTTAAACGGTATCACTACTACTTTC	7559	
Qy	241	ATTTCATCTCTCTGTTTCATTATGCTGACATATTTAAAGTCCGTTGTCAAATCTCTAAAGTT	300	
Db	7558	ATTTCATCTCTCTGTTTCATTATGCTGACATATTTAAAGTCCGTTGTCAAATCTCTAAAGTT	7499	
Qy	301	AGACATAAAGACAAAGACTAATCAATTTGTCAACAACAGGCTCGTGGAGTGAGCTATA	360	
Db	7498	AGACATAAAGACAAAGACTAATCAAATTTGTCAACAACAGGCTCGTGGAGTGAGCTATA	7439	
Qy	361	TTAATCTGGATTTTAAAGCATTTAAGAAAACATTTCTATAGTACTTAAAGCAAAATAAATAAT	420	
Db	7438	TTAATCTGGATTTTAAAGCATTTAAGAAAACATTTCTATAGTACTTAAAGCAAAATAAATAAT	7379	
Qy	421	TATATCAAAACACATATGCTTCACACTCGTCACTGTTAGTGTAGTGAATGATCTTACATC	480	
Db	7378	TATATCAAAACACATATGCTTCACACTCGTCACTGTTAGTGTAGTGAATGATCTTACATC	7319	
Qy	481	ATAAGAGGCGCATCAAAATCTTAAAAATAAGCATAAATGAAATTAATCATTTACAAATTTT	540	
Db	7318	ATAAGAGGCGCATCAAAATCTTAAAAATAAGCATAAATGAAATTAATCATTTACAAATTTT	7259	
Qy	541	ATTTTACTCAATAGAAAAATCGAAAGTATGATTTATCTAGCTGCCCAAACTTTTCGAAT	600	
Db	7258	ATTTTACTCAAATAAGAAAAATCGAAAGTATGATTTATCTAGCTGCCCAAACTTTTCGAAT	7199	
Qy	601	TTAATAATTTTACTCAGAAGAGACCGCATTTTAATCCCTTGACTTCTCATTTGCTCTATGGAA	660	
Db	7198	TTAATAATTTTACTCAGAAGAGACCGCATTTTAATCCCTTGACTTCTCATTTGCTCTATGGAA	7139	
Qy	661	AATGATTTAAGCAGTCAATAAAATCTTTTGCACATTTGTGGCAAGAACCAATAAATTCGAA	720	
Db	7138	AATGATTTAAGCAGTCAATAAAATCTTTTGCATTTGTGGCAAGAACCAATAAATTCGAA	7079	
Qy	721	GTCTAAATATGTAATCGTCCACACAGTGTATGAGTATCTCTAGTATATTTTTTTCTTTTCCAT	780	